CLAIMS:

1. A method of identifying a user-specified object contained in one or more images of a plurality of images, the method comprising defining regions of objects in said images, computing a vector in respect of each of said regions based on the appearance of the respective region, each said vector comprising a descriptor, vector quantizing said descriptors into clusters, storing said clusters as an index with the images in which they occur, defining regions of said user-specified object, computing a vector in respect of each of said regions based on the appearance of said regions, each said vector comprising a descriptor, and vector quantizing said descriptors into said clusters, searching said index and identifying which of said plurality of images contains said clusters so as to return the images containing said user-defined object.

2. A method according to claim 1, further comprising comparing the clusters relating to the objects contained in the images identified as containing an occurrence of said user-specified object with the one or more clusters relating to said user-specified object, and ranking said images identified as containing an occurrence of said user-specified object according to the similarity of the one or more clusters associated therewith to the cluster associated with said user-specified object.

3. A method according to claim 1, wherein at least two types of viewpoint covariant regions are defined in respect of each of said images.

4. A method according to claim 3, wherein a descriptor is computed in respect of each type of viewpoint covariant region.

5. A method according to claim 4, wherein one or more separate clusters are formed in respect of each type of viewpoint covariant region.

6. A method according to claim 3, wherein said at least two types of viewpoint covariant regions include Shape Adapted and Maximally Stable regions respectively.

7. A method according to claim 1, wherein said user-specified object is specified as a sub-part of an image.

8. A method according to claim 7, wherein identification of said user-specified object is performed by first vector quantizing the descriptor vectors in a sub-part of an image to precomputed cluster centers.

9.. A method according to claim 1, wherein the regions defined in each image are tracked through contiguous images and unstable regions are rejected.

10. A method according to claim 9, wherein an estimate of a descriptor for a track is computed from the descriptors throughout the track.

11. A method according to claim 1, wherein each image or portion thereof is represented by one or more cluster frequencies.

12. A method according to claim 11, wherein said cluster frequency is weighted.

13. A method according to claim 1, wherein a predetermined proportion of most frequently occurring clusters ·in said plurality of images are omitted from or suppressed in such index.

14. A method according to claim 1, wherein said index comprises an inverted file structure having an entry for each cluster which stores all occurrences of the same cluster in all of said plurality of images and possibly more precomputed information about each cluster occurrence such as for example its spatial neighbours in an image.

15. A method according to claim 1, including the step of ranking said images using local image spatial coherence or global relationships of said descriptor vectors.

16 A method of identifying a user-specified object contained in one or more image frames of a moving picture, the method comprising associating a plurality of different 'visual aspects' with each of a plurality of respective objects in said moving picture, retrieving the 'visual aspects' associated with said user- specified object, and matching said 'visual aspects' associated with said user-specified object with objects in said frames of said moving picture so as to identify instances of said user-specified object in said frames.

17. A method according to claim 16, wherein the 'visual aspects' associated with an object are obtained using one or more sequences or shots of a moving picture in which said object occurs.

18. A method according to claim 16, comprising tracking said object through a plurality of image frames in a sequence.

19. A method according to claim 18, comprising defining affine invariant regions of objects in said image frames and tracking one or more regions through a plurality of image frames in a sequence.

20. A method according to claim 19, comprising, in the event that a track terminates in an image frame of a sequence, propagating the track to either following or preceeding image frames in the sequence, so as to create a substantially continuous track throughout the image frames in the sequence.

21. A method according to claim 19 where tracked regions are grouped into objects according to their common motion using constraints arising from rigid or semi-rigid object motion.